

**Statement of Relevance for Japanese Laid-Open Patent No. 61-52698**

Laid-Open Date: March 15, 1986 (Shouwa 61)  
Entitled: A VOICE RECOGNITION APPARATUS  
Filing Serial No.: 59-174141  
Filing Date: August 23, 1984 (Shouwa 59)  
Inventor: Toshihiro Kimura  
Applicant: Hitachi Electronic Engineering Co. Ltd.

**Claim:**

A voice recognition apparatus comprising:

a voice analyzer means for serially analyzing inputted voice signals to obtain spectrum data of said voice signal;

a first storage means for recording standard patterns of voice spectra with respect to each of specific words depending on an application field;

a comparison computation means to which spectrum data analyzed in said voice analyzer means are inputted and for comparing said spectrum data with all of said standard patterns stored in said storage means to obtain similarities there-between; and

a determination means for selecting a standard pattern having the largest similarity and outputting voice recognition data corresponding to said selected standard pattern based on an output of said comparison computation means; said voice recognition apparatus further comprising:

a second storage means for storing each of standard patterns of voice spectra of corresponding predetermined meaningless words and other meaningless sounds often speeched by human beings; and

wherein said comparison computation means executes comparison said spectrum data analyzed in said voice analyzer means with all of said standard patterns of said meaningless words, etc. and said determination means outputs none of said voice recognition data when a standard pattern for said meaningless words, etc. is selected to be the largest similarity.

⑫ 公開特許公報(A)

昭61-52698

⑤ Int. Cl.

G 10 L 3/00

識別記号

GLA

庁内整理番号

7350-5D

8221-5D

④ 公開

昭和61年(1986)3月15日

審査請求 未請求 発明の数 1 (全5頁)

⑭ 発明の名称 音声認識装置

⑰ 特 願 昭59-174141

⑱ 出 願 昭59(1984)8月23日

⑲ 発 明 者 木 村 俊 宏 神奈川県足柄上郡中井町久所300番地 日立電子エンジニアリング株式会社内

⑳ 出 願 人 日立電子エンジニアリング株式会社 神奈川県足柄上郡中井町久所300番地

㉑ 代 理 人 弁理士 飯塚 義仁

明 細 書

発 明 の 名 称

音声認識装置

特許請求の範囲

入力された音声信号を逐次分析して該音声のスペクトルデータを得るための音声分析手段と、

利用分野に応じた特定の複数の語に関してその音声スペクトルの標準パターンを夫々記憶した第1の記憶手段と、

前記音声分析手段で分析されたスペクトルデータを入力し、該スペクトルデータと前記記憶手段に記憶された全ての標準パターンとを比較し、その類似度を求める比較演算手段と、

前記比較演算手段の出力に基づき、前記類似度の最も大きな標準パターンを選択し、該標準パターンに対応する音声認識データを出力するための判定手段と、

を具える音声認識装置において、

人間が発声しやすい無意味語及びその他の無意味音のうち所定の複数に関して各々の音声スペクトルの標準パターンを夫々記憶した第2の記憶手段を更に設け、

前記比較演算手段においては、前記音声分析手段で分析されたスペクトルデータと前記第2の記憶手段に記憶された全ての無意味語等の標準パターンとを更に比較し、

前記判定手段においては、前記無意味語等についての標準パターンを類似度の最も大きなものとして選択したときには、前記音声認識データを出力しないようにしたことを特徴とする音声認識装置。

発明の詳細な説明

〔産業上の利用分野〕

この発明は、人間の音声信号を入力し、該音声によって表現されたデータを認識するための装置に関し、特に「えー」、「えーと」等の無意味語の混入による誤認識を防ぐようにしたものに關す

る。

#### 〔従来の技術〕

マン・マシン・インターフェース改善の一つとして、単語ごとに区切って発音した音声認識対象とする音声認識装置が、近年実用化の域に達している。一例を示せば、銀行のキャッシュ・カード・サービスにおいては、カードの暗証番号等の照会を行なう際に預金者からの音声入力により該暗証番号を認識する装置が実用に供されている。

このような音声認識装置は、夫々の利用分野に応じて、あらかじめ種々の特定の単語（例えば数字等）についての音声スペクトルの標準パターンを保持している。該音声認識装置は、音声信号が入力されると、該音声信号をスペクトル分析してそのスペクトルデータを求め、該スペクトルデータと前記標準パターンの全てとの比較演算を逐次行なっていく。そして、全ての前記標準パターンの中から入力音声信号のスペクトルに対する類似度の最も高いものを選択し、その最も類似度の高い標準パターンに対応する所定の語を示すデータ

2, 6, 1」がこの預金者の暗証番号であると判断されてしまうのである。尚「えーと」のような無意味な言葉の他に、咳ばらいや特定の騒音が検出されたときにも、こうした現象が生じるおそれがある。

このように、従来の音声認識装置では、必要なデータを表現する音声以外の無意味な音声が入力されることにより、データの誤認識を犯してしまうという問題があった。

この発明は上述の点に鑑みてなされたもので、正確なデータの認識ができるようにした音声認識装置を提供しようとするものである。

#### 〔問題点を解決するための手段及び作用〕

この発明に係る音声認識装置は、該当する利用分野に応じた所定のデータについての音声スペクトルの標準パターンとは別途に、人間が発声しやすい無意味語若しくは咳ばらい、騒音等の無意味音についての音声スペクトルの標準パターンを記憶手段内に保持している。比較演算装置においては、取込まれたスペクトルは、前記データについ

て音声認識データ（入力音声から認識した語を示すデータ）として出力する。

#### 〔発明が解決しようとする問題点〕

しかしこのような従来の音声認識装置では、必要なデータを表現する音声だけでなく、それ以外の無意味な音声が入力されたときにも、上述の比較演算及び選択の結果、最も相対的な類似度の高い標準パターンが選択され、該パターンに対応して予め準備されている語がその音声によって表現された入力データであると誤認識されてしまう。

前述のキャッシュ・カード・サービスを例にとり、預金者が暗証番号を音声入力する際に「えーと、2, 6, 1, 5」と発音したとしよう。そうするとこの音声認識装置では、最初の「えーと」という音声についても上述の比較演算が行なわれ、この「えーと」という音声のスペクトルに対する類似度が相対的に最も高い標準パターンを有する数字（仮りに「3」であるとしよう）の選択が行なわれる。その結果、前記預金者の音声は「3, 2, 6, 1, 5」と認識され、最初の4桁の「3,

ての標準パターンだけでなく、この無意味語等についての全ての標準パターンとも逐次比較され、類似度が求められる。

判定手段は、スペクトル化された或る音声に対する類似度が最も大きなものとして、前記所定のデータについての標準パターンのひとつを選択したときは、該標準パターンによって表現されるデータを上位のコンピュータ等へ出力する。しかし類似度が最も大きなものとして前記無意味語についての標準パターンのひとつを選択したときは、該標準パターンによって表現される無意味語を無視し、出力しない。

これにより、所定のデータを音声入力する際に、その標準パターンが前記記憶手段に記憶されているところの無意味語を発声した場合でも、該無意味語は無視され、音声入力された正しいデータだけが正確に判定手段から出力される。

#### 〔実施例〕

以下、添付図面を参照してこの発明の一実施例を詳細に説明しよう。

第1図において、音声分析部1は、検出された音声の波形を逐次分析して該音声のスペクトルデータを得るためのものである。該分析部1では、例えば1秒間に800～1600バイトの割合で波形の分析が行なわれる。尚、分析部1では波形の分析を行なう前に該波形をディジタル信号に変換するのが通常であるが、該分析部1はアナログ波形のまま分析を行なうものであってもよい。

音声分析部1によって得られる音声のスペクトルデータを逐次取り入れるものとして、データ圧縮部2が設けられている。データ圧縮部2は、所定の時間間隔を単位として該スペクトルデータの間引き又は平均化等のデータ圧縮操作を行なうものである。前述のように毎秒800～1600バイトの割合で与えられる膨大な量のスペクトルデータは、これによって逐次演算処理を行なうのに適した量にまで減少させられる。尚この間引きまたは平均化は、音声の特徴を損なわないような範囲内で行ない、例えば該スペクトルデータを約10倍に圧縮する程度の時間間隔で行なうものとする。

についてのスペクトルの標準パターンが記憶される。

比較演算部3は、データ圧縮部2からスペクトルデータが与えられ始めると、これら全ての標準パターンをメモリ4及びメモリ5から順次読出して、該スペクトルデータと全ての標準パターンとの比較演算を逐次行なっていく。この比較演算の結果、比較演算部3からは、該スペクトルデータと各標準パターンとの類似度を夫々示す信号が逐次出力される。

比較演算部3の出力信号は、判定部6に逐次取入れられる。判定部6は、マイクロコンピュータまたは専用のハード装置から成っており、該信号に基づき、各入力音声のスペクトルデータに対する類似度の最も大きい標準パターンを選択するものである。その結果選択された標準パターンがメモリ4に記憶されているものであれば(すなわち意味のあるデータを表現するものであれば)、判定部6は、該標準パターンに対応する語(一語又は単語又は熟語等を含む)を示すデータを入力音

比較演算部3は、データ圧縮部2によって間引きまたは平均化されたスペクトルデータを逐次取り入れて、あらかじめ記憶された標準パターンと該スペクトルデータとの比較演算を行なうために設けられたものであり、マイクロコンピュータまたは専用のハード装置から成るものである。

この発明によれば、標準パターンを記憶するためのものとして、メモリ4及びメモリ5が設けられている。メモリ4は、利用分野に応じ、必要とされるデータを表わす音声のスペクトルの標準パターンを記憶するものである。例えば列車予約の場合であれば、数字、列車名、普通車かグリーン車かの区別、駅名等についての標準パターンがメモリ4に記憶される、といった具合である。他方、メモリ5は、該データを表わす音声以外の音声であってデータ入力時に発生しやすいものについてのスペクトルの標準パターンを記憶するためのものである。メモリ5には、例えば「えーと」、「えー」、「あー」のように人間が通常発声しやすい無意味語をはじめ、咳ばらい、使用場所に特有の騒音等

声の音声認識データとして図示しない上位のコンピュータ等へ出力する。また選択された標準パターンがメモリ5に記憶されているものであれば(すなわち無意味語に対応するものであれば)、判定部6は、これを見捨て、音声認識データの出力を行なわない。

ちなみに、比較演算部3及び判定部6がマイクロコンピュータから成るものであるとした場合に該マイクロコンピュータによって実行されるべきプログラムを略示すると、第2図のとおりである。

ステップ7及び8は比較演算部3の機能に相当するものであり、ステップ7では、データ圧縮部2から与えられたスペクトルデータとメモリ4から読出した真の標準パターンとの比較を逐次行ない、各々の類似度を求める。ステップ7が終了すると、ステップ8に進む。

ステップ8では、前記スペクトルデータとメモリ5から読出した無意味語等についての標準パターンとの比較を逐次行ない、各々の類似度を求める。ステップ8が終了すると、ステップ9に進む。

ステップ9～11は判定部6の機能に相当するものであり、ステップ9では、前記ステップ7及び8の比較結果に基づき前記スペクトルデータに対する類似度の最も大きい標準パターンを選択する。

続いてステップ10では、前記ステップ9で選択された標準パターンが、メモリ5に記憶された無意味語等についての標準パターンに属するものであるか否かが判断される。メモリ5内の標準パターンに属するものであれば、ステップ10でYESと判断し、そのままリターンする(従って上位コンピュータへの音声認識データの出力は行なわれない)。他方、メモリ5内の標準パターンに属さないものであれば(すなわち、メモリ4に記憶された真の標準パターンに属するものであれば)、ステップ10でNOと判断してステップ11に進み、ステップ9で選択された該標準パターンに対応する音声認識データが上位コンピュータに出力される。ステップ11を終了すると、リターンする。

ついで、メモリ5にその標準パターンが記憶されているところの無意味語のひとつである「えー」、「あー」を夫々最も類似度の大きいものとして選択し、これを無視する。従ってこの音声認識からは、列車名について「ひかり217号」、駅名について「新大阪」という正確なデータが出力されることになる。

このように、メモリ4にその標準パターンが記憶されているところのデータを音声入力する際に、メモリ5にその標準パターンが記憶されているところの無意味語を発声してしまった場合でも、この音声認識装置は、該無意味語等を無視し、音声入力された該データを正確に認識して外部に出力するのである。

なお、メモリ4及び5はハード的に別々のメモリであることを要せず、真のデータに対応する標準パターン及び無意味語等についての標準パターンの双方を一体のメモリハード回路に記憶したものであってもよい。

[ 発明の効果 ]

次に、具体例を交えてこの発明に係る音声認識装置の動作の一例を説明すると、以下のとおりである。列車予約において、予約者が列車名を音声入力する際「ひかり、えー、217号」、駅名を音声入力する際「あー、新大阪」と発音したとしよう。このとき該音声は、音声分析部1及びデータ圧縮部2によりスペクトル化され、間引きまたは平均化された後、比較演算部3において、メモリ4及びメモリ5内に記憶されている数字、列車名、駅名、無意味語等についての全ての標準パターンとの比較が行なわれ、その結果類似度を示す信号が該演算部3から判定部6に与えられる。このとき判定部6に与えられる。このとき判定部6は、「ひかり」、「217号」、「新大阪」という音声のスペクトルについては、メモリ4にその標準パターンが記憶されているところの駅名、列車名に関するデータのひとつである「ひかり」、「217号」、「新大阪」を夫々最も類似度の大きいものとして選択し、外部に出力する。他方「えー」、「あー」という音声のスペクトルに

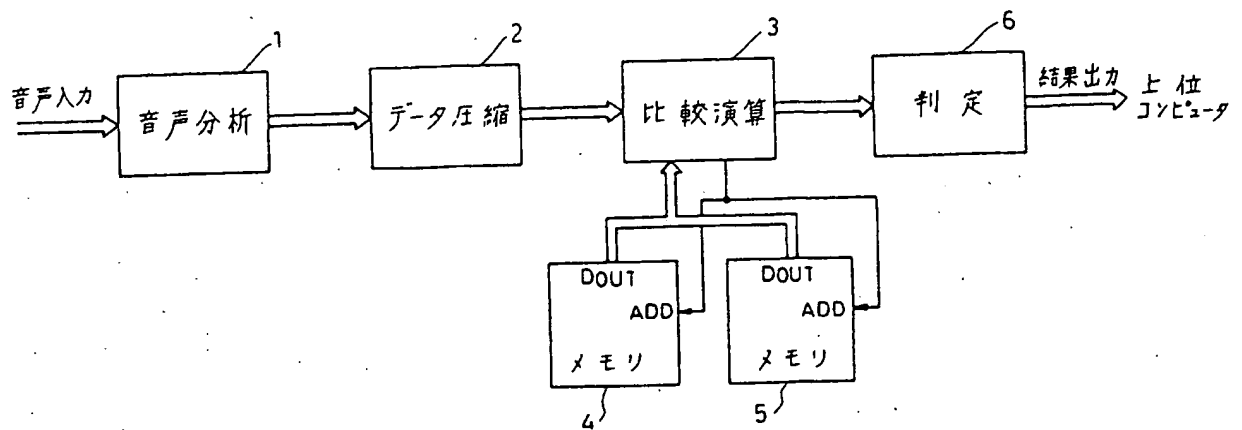
以上のとおりこの発明に係る音声認識装置によれば、データの音声入力中に無意味語等が入力された場合にも、該無意味語等の影響を受けることなく、正確に真の入力データの認識を行なうことができる。従ってまた、データの音声入力中に無意味語等が入力されても改めて該データの音声入力をやり直す必要がないので、全体としてのデータの入力時間を短縮することができる。これにより、音声認識装置を用いた各種サービス業務において、サービスの向上を図ることができる。

#### 図面の簡単な説明

第1図は、この発明に係る音声認識装置の一実施例を示す概略ブロック図、第2図は、該実施例における比較演算部及び判定部としてマイクロコンピュータを用いた場合に該マイクロコンピュータが実行すべきプログラムを略示するフローチャートである。

1…音声分析部、2…データ圧縮部、3…比較演算部、4、5…メモリ、6…判定部。

## 第 1 図



## 第 2 図

